

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Kensaku IMAI et al.

Serial No.: Not Yet Assigned

Group Art Unit:

Filed: Concurrently Herewith

Examiner:

For: **METHOD AND APPARATUS FOR AUTOMATICALLY REMOVING VECTOR UNIT
IN DNA BASE SEQUENCE**

STATEMENTS REGARDING SEQUENCE LISTINGS

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Attached is a paper copy of a "Sequence Listing."

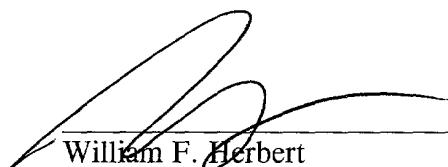
An acceptable computer readable form is on file with the parent application Serial No. 08/684,674, so no additional computer-readable copy is being submitted herewith.

I hereby state that the content of the attached paper copy and the computer readable copy on file of the Sequence Listing, submitted in accordance with 37 C.F.R. §1.821(c) and (e), respectively are the same.

I hereby state that the submission, filed in accordance with 37 C.F.R. §1.821(g), herein does not include new matter.

Respectfully submitted,

Date: 3/20/01



William F. Herbert
Registration No. 31,024

STAAS & HALSEY
700 Eleventh Street, N.W.
Suite 500
Washington, D.C. 20001

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: IMAI, Kensaku
KITAJIMA, Masato

(ii) TITLE OF INVENTION: METHOD AND APPARATUS FOR
AUTOMATICALLY
REMOVING VECTOR UNIT IN DNA BASE SEQUENCE

(iii) NUMBER OF SEQUENCES: 19

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Staas & Halsey
(B) STREET: 700 Eleventh Street, N.W., Suite 500
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: US
(F) ZIP: 20001

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/684,674
(B) FILING DATE: 22-JUL-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Herbert, William F.
(B) REGISTRATION NUMBER: 31,024
(C) REFERENCE/DOCKET NUMBER: 862.1335/WFH

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 2024341500
(B) TELEFAX: 2024341501

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AAGCTTGCAT GCCTGCAGGT CGACTCTAGA GGATCCCCGG GTACCGAGCT CGAATTC

57

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGCACTTGAA CGCATGCT

18

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGCACTTGAA CGCTGCT

17

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TGCACTTGAC GCATGCT

17

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGCACTTGAC GCATGCT

17

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TGCCTTGAAC GCATGCT

17

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2686
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA

60

CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG

120

TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC

180

ACCATATGCG	GTGTGAAATA	CCGCACAGAT	GCGTAAGGAG	AAAATACCGC	ATCAGGCGCC	240
ATTGCCATT	CAGGCTGCGC	AACTGTTGGG	AAGGGCGATC	GGTGCAGGGCC	TCTTCGCTAT	300
TACGCCAGCT	GGCGAAAGGG	GGATGTGCTG	CAAGGCGATT	AAGTTGGGTA	ACGCCAGGGT	360
TTTCCCAGTC	ACGACGTTGT	AAAACGACGG	CCAGTGCCAA	GCTTGCATGC	CTGCAGGTCTG	420
ACTCTAGAGG	ATCCCCGGGT	ACCGAGCTCG	AATTGTAAT	CATGGTCATA	GCTGTTCCCT	480
GTGTGAAATT	GTTATCCGCT	CACAATTCCA	CACAACATAC	GAGCCGGAAG	CATAAAGTGT	540
AAAGCCTGGG	GTGCCTAATG	AGTGAGCTAA	CTCACATTAA	TTGCAGTTGCG	CTCACTGCC	600
GCTTCCAGT	CGGGAAACCT	GTCGTGCCAG	CTGCATTAAT	GAATCGGCCA	ACGCGCGGGG	660
AGAGGCGGTT	TGCGTATTGG	GCGCTCTTCC	GCTTCCTCGC	TCACTGACTC	GCTGCGCTCG	720
GTCGTTCGGC	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG	CGGTAATACG	GTTATCCACA	780
GAATCAGGGG	ATAACGCAGG	AAAGAACATG	TGAGCAAAAG	GCCAGCAAA	GGCCAGGAAC	840
CGTAAAAGG	CCGCAGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	900
AAAAATCGAC	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GAATATAAAG	ATACCAGGCG	960
TTTCCCCCTG	GAAGCTCCCT	CGTGCCTCT	CCTGTTCCGA	CCCTGCCGCT	TACCGGATAC	1020
CTGTCGCGCT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTCCTC	AAAGCTCACG	CTGTAGGTAT	1080
CTCAGTTCGG	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCAG	1140
CCCGACCGCT	GCGCCTTATC	CGGTAACATAT	CGTCTTGAGT	CCAACCCGGT	AAGACACGAC	1200
TTATGCCAC	TGGCAGCAGC	CACTGGTAAC	AGGATTAGCA	GAGCGAGGTA	TGTAGGCGGT	1260
GCTACAGAGT	TCTTGAAGTG	GTGGCCTAAC	TACGGCTACA	CTAGAAGAAC	AGTATTGTT	1320
ATCTGCGCTC	TGCTGAAGCC	AGTTACCTTC	GGAAAAAGAG	TTGGTAGCTC	TTGATCCGGC	1380
AAACAAACCA	CCGCTGGTAG	CGGTGGTTTT	TTTGGTTGCA	AGCAGCAGAT	TACGCGCAGA	1440
AAAAAAGGAT	CTCAAGAAGA	TCCTTGATC	TTTCTACGG	GGTCTGACGC	TCAGTGGAAC	1500
GAAAACTCAC	GTAAAGGGAT	TTTGGTCATG	AGATTATCAA	AAAGGATCTT	CACCTAGATC	1560
CTTTAAATT	AAAAATGAAG	TTTTAAATCA	ATCTAAAGTA	TATATGAGTA	AACTGGTCT	1620
GACAGTTACC	AATGCTTAAT	CAGTGAGGCA	CCTATCTCAG	CGATCTGTCT	ATTCGTTCA	1680
TCCATAGTTG	CCTGACTCCC	CGTCGTGTAG	ATAACTACGA	TACGGGAGGG	CTTACCATCT	1740
GGCCCCAGTG	CTGCAATGAT	ACCGCGAGAC	CCACGCTCAC	CGGCTCCAGA	TTTATCAGCA	1800

ATAAACCAGC CAGCCGGAAG GGCGAGCGC AGAAGTGGTC CTGCAACTT ATCCGCCTCC	1860
ATCCAGTCTA TTAATTGTTG CCGGGAAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTG	1920
CGCAACGTTG TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGT TGGTATGGCT	1980
TCATTCAGCT CCGGTTCCCA ACGATCAAGG CGAGTTACAT GATCCCCAT GTTGTGCAAA	2040
AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTAGAA GTAAGTTGGC CGCAGTGTAA	2100
TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC	2160
TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG	2220
AGTTGCTCTT GCCCGGCGTC AATACGGGAT AATACCGCGC CACATAGCAG AACTTTAAAA	2280
GTGTCATCA TTGGAAAACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT ACCGCTGTTG	2340
AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC	2400
ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAAATAAGG	2460
GCGACACCGA AATGTTGAAT ACTCATACTC TTCCTTTTC AATATTATTG AAGCATTAT	2520
CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA TAAACAAATA	2580
GGGGTTCCGC GCACATTTCC CCGAAAAGTG CCACCTGACG TCTAAGAAC CATTATTATC	2640
ATGACATTAA CCTATAAAAA TAGGCGTATC ACGAGGCCCT TTCGTC	2686

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 66
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTGCCAAGCT TGCATGCCTG CAGGTGACT CTAGAGGATC CCCGGTACCG AGCTCGAATT	60
CGTAAT	66

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAGCTT

6

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCATGC

6

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTGCAG

6

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGTAC

6

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCTAGA

6

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GTCGAC

6

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCGAC

6

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCCGGG

6

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GAATTC

6

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCCGGG

6

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAATTC

6